CHAPTER 5

WATER QUALITY DATA COLLECTION AND ANALYSIS

Section I. Introduction

5-1. Purpose. This chapter provides guidance on collecting water quality data, database management, and data presentation. It is intended for scientists and engineers responsible for compiling existing data from non-Corps sources, establishing Corps data collection programs, and analyzing data for District or Division water quality management programs.

5-2. Overview. Since results from data-gathering efforts from non-Corps sources and Corps water quality sampling programs are used to evaluate both existing reservoir water quality conditions and the attainment of reservoir water quality objectives, these assessment programs must be well conceived and scientifically sound and must provide representative water quality data.

a. Objectives of the Study. A clear, concise statement of the study objectives is essential (Chapter 3). Without a statement of objectives, it is unlikely that the right questions will be addressed, the appropriate data collected, or the proper analyses performed. Sampling programs are expensive, so it is important to ensure that the data collected are useful. As data are analyzed and the water quality concerns and processes are better understood, water quality objectives may be clearly defined and the sampling program modified to address areas of uncertainty or concern.

b. Sampling Objectives. The sampling program objectives must be specifically defined and documented to ensure successful implementation and completion of the program. To a large extent, defining the objectives will determine the data needs in the program. The importance of the interrelated factors such as objectives, system characteristics, and degree of precision should also be considered (Ref. 64).

c. Population to Be Characterized. The term population is defined as the assemblage from which the sample is taken. There is the overall or general population and a sample population. The general population represents the entire set of measurements about which inferences or conclusions are to be made (e.g., the entire reservoir or the entire algal assemblage). The sample population (i.e., measurement of that constituent in the water quality samples) represents a subset of measurements taken from the general population. The subset is used to gain information and make inferences about the overall population. It is important that the sampled population be representative of the general population. The purpose of sampling design is to ensure the limited number of samples collected (e.g., 25-50 one-liter samples per sampling trip) provide adequate information about the overall population characteristics (e.g., total amount of that constituent in a reservoir containing about 10^{11} liters of water). The initial sample population, then, must not become altered or redefined during the sampling program.

d. Characteristics to be Measured. After the general population has been clearly defined, all relevant population characteristics should be identified and subdivided into essential or ancillary categories. The key to reducing the uncertainty surrounding project water quality is to collect samples pertaining to essential population characteristics. Sampling effort on ancillary characteristics that indirectly affect or are of secondary importance in attaining the study objectives should be minimized. Identifying these characteristics can be a difficult task but is an important requirement to ensure appropriate use of resources (see Chapter 2 for additional information).

e. Degree of Precision. All measurements have inherent errors or uncertainty because only part of the population has been measured. The degree of uncertainty can be reduced by taking more samples and using more precise techniques. However, budgetary constraints ultimately limit these approaches. Precision and cost are essential and interrelated elements in all sampling programs since precision influences cost. If the desired precision results in excessive cost, the precision of some or all variables may have to be reduced. Initial estimates of desired precision should reflect the analytical precision of the individual constituent. Estimates of analytical precision have been determined by the American Public Health Association (APHA) (Ref. 48) and the US Geological Survey (USGS) (Ref. 46).

f. Selection Criteria. The criteria established for selecting a sampling design represent a compromise between the desired precision and budget constraints. With any sampling design, the sample size is dependent on precision, cost, and variability of the estimator used to describe a population characteristic. When the sample size for each design has been determined, relative costs and time involved for each design can be computed and decisions relative to the acceptable sampling design can be made.

g. Use of Existing Data. As part of determining the data needs for an assessment, a screening of existing data should be conducted (see Chapter 4, Section I). Using available data can save effort and money; however, an appropriate level of confidence must be built into these data (e.g., Is the source known? Were the methods of collection and analysis used compatible with the level of detail and quality identified in the objectives of the proposed assessment?). In some cases, existing data may be used in the assessment or they may provide an "order of magnitude" reference from which to begin data collection for the assessment. In instances where little is known about the data, it is better to collect data specifically for the assessment.

Section II. Field Data Collection

5-3. Principles. Green (Ref. 69) cites ten basic principles of sampling design and statistical analysis for environmental studies that should be considered in developing a reservoir water quality assessment study. These principles are:

a. Clearly identify the objectives. An investigator must be able to describe clearly to someone else the purpose of the investigation. The results can only be as coherent and as comprehensible as the initial definition of the problem.

b. Make sure the investigator takes replicate samples to get an estimate of the variability and uncertainty in the sampling program.

c. Make sure the investigator has a random sampling program.

d. To test whether a condition has an effect, collect samples both where the condition is present and where the condition is absent. It is also important to collect samples where other factors are similar. An effect can only be demonstrated by comparison with a control.

e. Carry out preliminary sampling to provide a basis for evaluation of sampling design and statistical analysis options. Skipping this step in an effort to save time usually is unproductive.

f. Verify that the sampling device or method is sampling the characteristics, area, variables, or organisms that are needed, with equal and adequate efficiency over the entire range of sampling conditions to be encountered. Variations in sampling efficiency of anaerobic versus aerobic samples from area to area can, for example, bias comparisons.

g. If the area to be sampled has a large-scale environmental pattern, break the area up into relatively homogeneous subareas and allocate samples to each in proportion to the size of the subarea.

h. Verify that the sample unit size is appropriate to the size, densities, and spatial distributions of the variable that is being sampled. Also, estimate the number of replicate samples required to obtain the precision wanted.

i. Test the data to determine whether the error variation is homogeneous, normally distributed, and independent of the mean. If it is not, as may be the case for most field data, then appropriately transform the data, use a distribution-free (nonparametric) procedure, use an appropriate sequential sampling design, or test against simulated data.

j. Having chosen the best statistical method to test one's hypothesis, it is important to stick with the result. An unexpected or undesired result is not a valid reason for rejecting the method and seeking a "better" one.

5-4. Sampling Designs. The study objectives, specified precision, and cost usually dictate which sampling design is implemented. Since the purpose of the sampling design is to characterize some aspects of reservoir water quality, characteristics of the general population in the reservoir must be considered. Longitudinal gradients, tributary sources, and other patterns can increase the variability in the data. The sampling design should consider these sources of variability at the beginning of the sampling program so procedures can be incorporated to minimize their impact during the later data analysis and interpretation phases. This can be accomplished through various types of random sampling.

a. Simple Random Sampling. Simple random sampling is a method of selecting n sampling units out of the total N units so that every sampling unit has an equal chance of being selected. With reservoir water quality sampling programs, the sampling units usually correspond to sampling stations or locations. Stations can be selected by superimposing a grid system on the reservoir water surface and selecting locations at random. Sampling depths should be selected at random if vertical resolution is desired. The assumption of homogeneity among sampling units is critical in the simple random sampling scheme. If the population cannot be divided into N homogeneous sampling units, then simple random sampling should not be used. However, if the assumption is valid (i.e., the reservoir is relatively homogeneous), this approach provides an efficient, cost-effective procedure. Assessing the variability in surface chlorophyll concentrations may use a simple random sampling approach.

(1) Estimation of sample size. The number of samples to be collected depends on the variability of the population characteristics that are being estimated and the desired precision of the estimate. Estimates of variability can be obtained by reviewing existing data or by conducting surveys or reconnaissance studies on the reservoir. The general formula for sample size determination is:

$$n = \frac{t^2 s^2}{d^2}$$

where

n = number of samples

t = appropriate value from Student's t distribution

s^2 = sample variance

d = desired precision about the mean

Since Student's t-value varies as a function of n , a t-value for 30 degrees of freedom can be used to initialize the procedure. The formula, then, can be evaluated iteratively, substituting the appropriate t-value for each predicted n until the iterative procedure converges (see Appendix C).

(2) Sample size determination for multiple characteristics. If sampling objectives are to obtain information on several water quality variables,

sample size determination becomes more tedious to achieve the desired precision for each variable. Sample size estimates must be determined for each constituent and the maximum estimates chosen as the sample size for the sampling plan. For those variables that require significantly fewer samples to achieve the desired precision, a randomization scheme can be constructed to subsample the sampling units. Another approach is to estimate the sample size only for the most critical variables. Although this does not guarantee the desired precision for all constituents, precision for the most critical constituents can be achieved; precision for the other variables can be estimated and will be available for future use and analyses.

b. Stratified Random Sampling. With stratified random sampling, the population of N units is divided into subpopulations of N_1, $N_2 \ldots N_L$ units, respectively. The subpopulations are nonoverlapping and together comprise the entire population, so that

$$N_1 + N_2 + \ldots + N_L = N$$

These subpopulations are called strata. To obtain maximum benefit from stratification, the values of N_j ($j = 1, 2 \ldots L$) must be known. In a typical water quality sampling program, strata may be the epilimnion, metalimnion, and hypolimnion; or the headwater, main pool, and near-dam locations; or a combination of these or other designations. Once the strata have been determined, samples are independently drawn from each stratum. The sample sizes within each stratum are denoted by n_1, $n_2 \ldots n_L$.

(1) Estimation of sample size. Stratification of the population produces a highly efficient sample allocation scheme. A general formula for sample size distribution is

$$n = \frac{(\Sigma W_i S_i)^2}{d^2/t^2}$$

where

n = total number of samples

W_i = weighting factor for stratum i (e.g., ratio of volume of stratum to total volume or surface area of stratum to total surface area)

S_i = standard deviation of samples in stratum i

d = desired precision

t = appropriate Student's t-value

Sample size within a stratum can be determined by

$$\frac{n_i}{n} = \frac{W_i S_i}{\Sigma(W_i S_i)}$$

Stratified random sampling in reservoirs usually requires fewer samples than random sampling to obtain the same precision (Appendix C).

(2) Optimum allocation. The sample size allocation scheme for each individual stratum can incorporate cost, if sampling costs vary over strata. The objective is to allocate samples to minimize costs. The simplest cost function is of the form

$$C = C_o + \Sigma C_i N_i$$

where

C = total cost

C_o = fixed cost (e.g., travel, per diem, etc., per sampling trip)

C_i = cost per sample within stratum i

N_i = the number of samples within stratum i

With this as the underlying cost function, the n_i's from the stratified sampling formula can be determined by

$$n_i = n \left[\frac{W_i S_i / C_i}{\Sigma(W_i S_i / C_i)} \right]$$

An example of a cost-precision-probability matrix is shown in Table 5-1. A similar table can be developed for other variables and used to optimize the sampling scheme to retain precision on critical constituents and reduce precision or probability requirements on desired constituents to satisfy funding constraints.

(3) Fixed sample numbers. The sampling formula can also be used to assess the loss of precision if fixed numbers of samples are collected at each station. It is unrealistic to expect a field crew to collect nine phosphorus samples, three turbidity samples, and twelve chlorophyll samples at station 3 (Table 5-1) with a variable number of samples at other stations. If it is determined that six samples will be collected from the epilimnion, four samples from the metalimnion, and eight samples from the hypolimnion at every station, the precision for each constituent can be determined by rearranging the sampling formula as

TABLE 5-1

Example of a Sampling Matrix to Optimize Sample Numbers, Precision, and Cost (after Ref. 104)

		Station 3						Station 10 or 12						Station 14 or 15					
Total Phosphorus	Unit Cost	$13						$13						$13					
	Mean[1]	9 μg P/ℓ						20 μg P/ℓ						34 μg P/ℓ					
	Precision[2]	±50%			±100%			±25%			±50%			±15%			±30%		
	Probability	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%
	Sample No.	12	9	5	4	4	3	18	13	8	6	5	3	32	23	14	10	7	5
	Total Cost[3]	156	117	65	52	52	39	234	169	104	78	65	39	416	299	182	130	91	65
Turbidity	Unit Cost	$3						$3						$3					
	Mean	1.3 NTU's						5.3 NTU's						5.3 NTU's					
	Precision	±50%			±100%			±20%			±40%			±20%			±40%		
	Probability	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%
	Sample No.	13	10	6	5	4	3	8	6	4	4	2	1	8	6	4	4	2	1
	Total Cost	39	39	18	15	12	9	24	18	12	12	6	3	24	18	12	12	6	3
Chlorophyll	Unit Cost	$20						$20						$20					
	Mean	2 μg/ℓ						7 μg/ℓ						11 μg/ℓ					
	Precision	±25%			±50%			±25%			±50%			±15%			±30%		
	Probability	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%	95%	90%	80%
	Sample No.	19	13	9	6	5	3	13	10	6	5	4	3	21	15	9	7	5	4
	Total Cost	380	260	180	120	100	60	260	200	120	100	80	60	420	300	180	140	100	80

an values are those expected based on three sampling dates.
ent precision based on approximate levels of analytical precision for each test or requirements of the study.
cost calculated as product of unit cost and sample number.

$$d = \frac{ts}{\sqrt{n}}$$

for the random sampling formula, and for the stratified random sampling formula

$$d = \frac{(\Sigma W_i S_i)t}{\sqrt{n}}$$

While the desired precision may not be obtained for every water quality constituent, the precision and uncertainty associated with each constituent can be determined and factored into decisions related to water quality.

c. Other Techniques. The two sampling approaches discussed in the preceding paragraphs are the most commonly used. However, other sampling schemes are available and described in the technical literature. If the two approaches discussed above are not acceptable, then such schemes as systematic sampling, one-stage sampling, two-stage sampling, etc., could be used. More complete descriptions of these and other techniques can be found in Snedecor and Cochran (Ref. 99) or Winer (Ref. 111).

5-5. Field Sampling and Analysis. Field sampling procedures, methodology, and analyses have been discussed in detail by APHA (Ref. 48), EPA (Ref. 39), Likens and Wetzel (Ref. 86), and USGS (Ref. 46). Several excellent points on field sampling are made by Kittrell and West (Ref. 84), although stream sampling is emphasized. This paragraph follows the discussion in the National Handbook of Recommended Methods for Water-Data Acquisition (Ref. 46). On field sampling activities, major topics include obtaining representative samples, maintaining quality control and assurance in the field, and selecting field personnel.

a. Representative Samples. A number of factors are important in obtaining representative water quality samples. These factors include sampling objectives; station location, depth, and frequency; equipment; sample variables; sample handling and preservation; and sample identification.

(1) Sampling objectives. The sampling program should be dictated by the sampling objectives. Monitoring programs, water quality surveys, intensive sampling, or regulatory sampling may have objectives that require different field sampling designs and procedures. For example, water quality surveys may have lower precision requirements and a more restricted budget than intensive sampling programs, and therefore require different sampling designs. Procedures for maintaining the chain of custody, for example, are not critical in most monitoring programs but are extremely critical in addressing legal questions. Any change in objectives must be accompanied by a review of the sampling program and can necessitate a change in the sampling approach.

(2) Sampling locations. Sampling station locations are influenced by the sampling objectives, hydrology, reservoir and hydraulic outlet geometry, point and nonpoint sources, accessibility, available equipment and facilities, and personnel. Flow estimates must accompany all tributary and reservoir release water quality samples. Tributary sampling stations, therefore, may be selected to correspond with gaging stations. If this is not possible, flow must be measured at the time of sampling. Sampling stations should be located in representative areas of the reservoir and are dictated by the sampling program objectives. For example, estimating mean or average conditions for the reservoir may result in samples collected proportional to reservoir volume. The greatest number of samples would be collected from the mixed layer since this represents the greatest volume in the reservoir, with a progressive decrease in the number of samples collected with depth. Longitudinal and lateral variability, as well as vertical variations in many water quality constituents, may be considered in determining station locations. At a minimum, samples should be collected in the inflow, the outflow, and at a representative station in the pool.

(a) Longitudinal variation. Many reservoirs have areas with distinct water quality conditions, such as the headwater, zone of transition, or lacustrine area. The extent of these areas should be identified before locating sampling stations. The zone of transition is a function of the plunge point depth, so the minimum and maximum depths for the plunge point can be computed to delineate the headwater and lacustrine areas. Formulations for estimating the plunge point depth are available (Ref. 12). Annual high and low quartile flows (i.e., 75 and 25 percent of median annual flow) can be used to predict the maximum and minimum plunge point depths, respectively, in the reservoir. Stations located upstream of the minimum plunge point depth should be in the headwater area, stations located downstream from the maximum plunge point depth should be in the lacustrine, while stations in between the minimum and maximum plunge point depths can be used to characterize the zone of transition. This longitudinal variability also may occur from the headwater of a cove to its confluence with the main body of the reservoir.

(b) Lateral variability. Tributary inflows tend to follow the old channel or thalweg through the reservoir both as underflows and interflows. The zone of conveyance for flow and associated constituents may not extend across the reservoir. This can result in lateral differences across the reservoir. Initial sampling efforts should investigate potential lateral variability by sampling over the thalweg and at alternative locations across the reservoir. This can be evaluated initially by specific conductivity measured laterally across the reservoir. If conductivity is relatively constant, most dissolved constituents probably are also. Transmissometer readings across the reservoir may indicate particulate constituent variability. Established sediment survey transects provide permanent reference points, transect geometry, and representative reservoir areas and should be used initially.

(c) Bridges. Bridges are typically selected as sampling sites for reservoirs because of accessibility and convenience. Bridges crossing the reservoir in the headwater area may be satisfactory sites since vertical stratification is usually minimal while boat access may be limited. Bridges, however, are generally located or constructed at reservoir constrictions. These localized constrictions may result in localized velocity increases that disrupt stratification patterns and result in an altered water quality regime around the bridge. Therefore, sampling from bridges may not provide representative data for reservoir areas where vertical stratification occurs. Bridges or other constrictions also may effectively isolate various parts of the reservoir. Backwater effects and altered sedimentation regimes may result in different water quality in these areas and may require sampling stations to characterize water quality in this area.

(3) Sample depths. Sampling objectives, reservoir geometry, hydraulic outlet design, hydrology, stratification patterns, and reservoir operation all influence selection of appropriate sampling depths. A fixed-depth sampling approach generally is adequate if sufficient samples are collected to characterize water quality throughout the water column. Fixed sampling depths from the reservoir water surface represent the most common approach and permit seasonal and year-to-year water quality comparisons at and among stations in monitoring programs. Using the reservoir water surface as the reference point also permits comparisons, even with large variations in the water surface elevation. It is recommended that sampling depths be selected to characterize the epilimnetic, metalimnetic, and hypolimnetic stratification zones while integrated samples may be considered to characterize the upper mixed layer and reduce the number of epilimnetic samples. Since development and deepening of the thermocline during the stratification period changes the mixed layer depth, integrated samples may be collected based on the average mixed layer depth for this period. At least three samples--surface, middepth, and bottom--should be collected in reservoirs that generally remain well mixed, since these reservoirs may intermittently stratify. Water quality conditions may change drastically during these intermittently stratified periods. Middepth and bottom samples are particularly important in these reservoirs during ice cover when oxygen concentrations may be depleted. Anoxic conditions during winter are typically initiated at the sediment/water interface.

(4) Sampling frequency. Sampling frequency or sampling times and dates are critical in obtaining representative reservoir water quality data. Fixed-interval sampling may miss the important hydrologic and limnological events occurring in a reservoir. For example, a typical monthly sampling program (e.g., every 30 days) in the Caddo River tributary to DeGray Lake during 1977 would not have incorporated storm flow in any of the samples (Figure 5-1). The majority of nutrient, suspended sediment, and other constituent loading, however, may occur during elevated flow periods. The sampling intervals should incorporate the important hydrological and limnological events affecting reservoir water quality (Table 5-2). The same total number of samples may be collected in either a fixed-interval, monthly sampling program or variable-interval sampling program, but more information per sample and more insight into reservoir water quality can be obtained from variable-interval sampling.

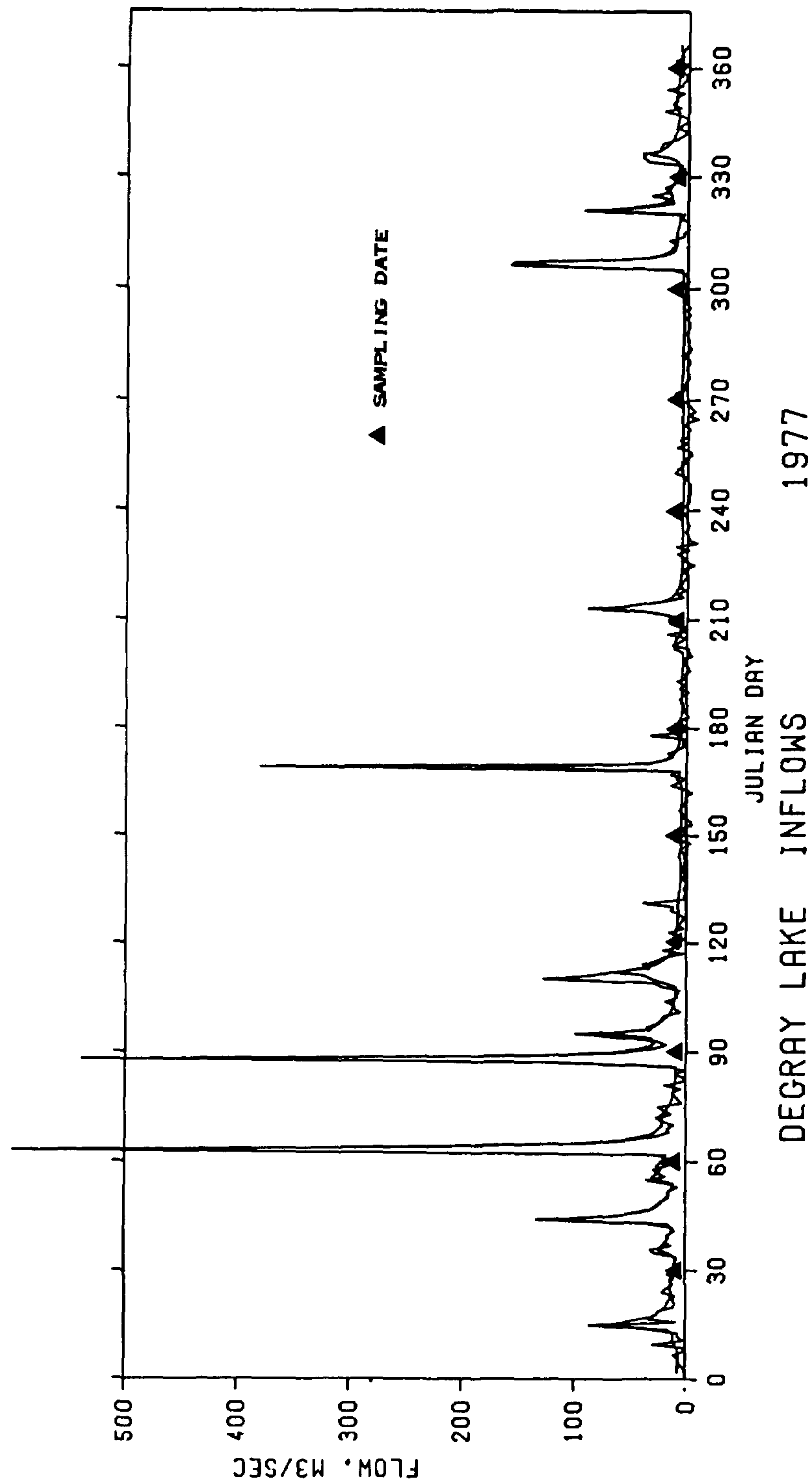


Figure 5-1. Potential bias in sampling program using a fixed interval (e.g., 30-day) sampling period

TABLE 5-2

Example of Sampling Intervals Corresponding with Hydrologic and Limnological Periods

Date Sampled	Event
Mid-March	Isothermal or late winter ice period
Mid-late April/ mid-May	Elevated flow; early stratification
Early-late June/ mid-July	Increased biological activity and public use
Early-late August	Strong stratification, low runoff
Late September/ mid-October	Anoxic conditions, plankton blooms, low flow
Mid-late November	Thermocline deepening and turnover; isothermal, initial winter conditions

Similar information can be obtained by combining special-interval or event sampling with a fixed-interval sampling program.

(5) Sampling equipment. Field gear typically used in water quality sampling is described in APHA (Ref. 48), USGS (Ref. 46), and Likens and Wetzel (Ref. 86) and is discussed and demonstrated through training courses offered by the EPA, USGS, and other Federal agencies. The use of sampling equipment, like any other analytical procedures or techniques, requires that assumptions and limitations be considered. Metal water samplers, for example, should not be used to obtain water samples for metal analyses or primary productivity measurements. Plastic samplers and bottles may interfere with organic and certain trace metal analyses. Two grab samplers used extensively in many field sampling programs are the Van Dorn and Kemmerer samplers. These samplers can collect water samples at any given depth or point in the water column. Samples also may be collected by pumping water from a given depth to the surface for collection. Pumped samples are advantageous when large numbers of samples are to be collected, the reservoir is not exceptionally deep (e.g., <150 feet), or anoxic samples are to be collected. When collecting pumped samples, the hoses must be allowed to clear and flush the water from a previous depth before filling the sample containers. In situ measurements can be made using sensor probes and digital or analog readout or recording devices. These instruments can measure temperature, DO, pH, orthophosphate, specific conductance, several specific cations and anions, and light penetration. While grab samplers, pump systems, and in situ probes all may be required to obtain representative water quality data, their use must be based on a clear understanding of the data needs, data use in water quality management, and characteristics of the reservoir. Regardless of the standardization or accepted use of sampling gear, all field sampling equipment, from a thermistor to a Kemmerer sampler, have idiosyncrasies. User experience and familiarity are critical in obtaining representative water quality samples or measurements.

(6) Sample variables. The water quality variables incorporated in the sampling program are a function of the project purposes, sampling objectives, applicable water quality standards or criteria, facilities and equipment, personnel, and funding constraints. Some water quality constituents typically measured in reservoir water quality sampling programs are listed in Table 5-3. Incorporating variables that can be measured in situ or surrogate variables can provide some sampling economies. It must be recognized, however, that surrogate variables provide only inferential information and not direct estimates for the primary variable of interest.

(a) In situ variables. Water temperature represents the most common in situ variable measured. Incorporating the capability to measure DO, specific conductivity, and pH in the same instrument adds little to the overall cost of the sampling program but can add significantly more information about reservoir water quality. The time required to measure these additional constituents also is insignificant. In situ data can be collected using remote sensing techniques or continuous monitoring. Remote sensing includes the

TABLE 5-3

Typical Water Quality Variables Measured in Reservoirs and the Sample Handling and Preservation Requirements

Determination	Container[1]	Min. Sample Size, ml	Preservation	Max. Storage Time Recommended/Regulatory
Acidity	P,G(B)	100	Refrigerate[2]	24 hr/14 days
Alkalinity	P,G	200	Refrigerate	24 hr/14 days
Carbon, organic total	G	100	Analyze immediately; or refrigerate and add H_2SO_4 to pH <2	7 days/28 days
Carbon dioxide	P,G	100	Analyze immediately	
Chlorophyll	P,G	500	30 days in dark; freeze	30 days/--
Color	P,G	500	Refrigerate	48 hr/48 hr
Conductivity	P,G	500	Refrigerate	28 days/28 days
Hardness	P,G	100	Add HNO_3 to pH <2	6 months/6 months
Metals, general	P(A),G(A)	-	For dissolved metals filter immediately, add HNO_3 to pH <2	6 months/6 months
Nitrogen:				
Ammonia	P,G	500	Analyze as soon as possible or add H_2SO_4 to pH <2; refrigerate	7 days/28 days
Nitrate	P,G	100	Add H_2SO_4 to pH <2	48 hr/48 hr
Nitrate, nitrite	P,G	200	Analyze as soon as possible or refrigerate; or freeze at -20° C	none/28 days
Organic, Kjeldahl	P,G	500	Refrigerate; add H_2SO_4 to pH <2	7 days/28 days
Oxygen dissolved:	G.BOD bottle	300	Analyze immediately	0.5 hr/1 hr
Electrode			Titration may be delayed after acidification	8 hr/8 hr
pH	P,G	-	Analyze immediately	2 hr/2 hr
Phosphate	G(A)	100	For dissolved phosphate forms filter immediately, refrigerate; freeze at 10° C	48 hr/48 hr
Residue	P,G	-	Refrigerate	7 days/7-14 days
Salinity	G.wax seal	240	Analyze immediately or use wax seal	6 months/--
Silica	P	-	Refrigerate, do not freeze	28 days/28 days
Sulfate	P,G	-	Refrigerate	28 days/28 days
Sulfide	P,G	100	Refrigerate; add 4 drops 2 N zinc acetate/100 ml	28 days/28 days
Temperature	P,G	-	Analyze immediately	
Turbidity	P,G	-	Analyze same day; store in dark up to 24 hr	24 hr/48 hr

SOURCE: Ref. 48.

[1] P = plastic (polyethylene or equivalent); G = glass; (A) = acid rinsed; (B) = borosilicate.
[2] Refrigeration = storage at 4° C.

use of both aerial and satellite imagery to collect water quality information. Engineer Pamphlet 70-1-1 fully explains the applications of remote sensing and should be referenced for specific information. The field sampling program can provide the data required to correlate spectral densities with water quality. Remote sensing can provide information on surface variability in water quality constituents, reservoir circulation patterns, inflow mixing processes, and other surface phenomena. An economical approach to obtaining continuous or time-series data for selected water quality constituents is to employ continuous water quality monitors. Continuous monitors may collect single- or multiple-constituent measurements in situ with a probe or sensing unit. Water temperature is the water quality constituent typically monitored, with DO, pH, and specific conductance also commonly measured. Continuous monitors have generally been used to monitor inflow and release water quality constituents since vertical gradients are weak and the sensing unit for in situ measurements or pump intake can be located at a representative point in the stream. Continuously monitoring reservoir water quality may require multiple sensors or pump intakes located vertically throughout the water column. For some applications, this can be the most economical means of obtaining continuous or time-series data. Presently, however, continuous monitors are applicable primarily for monitoring certain inflow and release water quality characteristics. Continuous monitoring is not synonymous with maintenance-free data collection. Continuous monitors require routine maintenance, generally on a weekly or biweekly basis, to ensure the instruments are functioning and remaining within calibration tolerances. Continuous monitors, units of measure, definitions, and considerations have been described in detail by the USGS (Ref. 46). The US Environmental Protection Agency (Ref. 40) and Schofield (Ref. 97) have reviewed automatic samplers and sampler design. These documents should be reviewed for additional information on continuous monitors. Data reduction and data management should receive careful consideration for monitoring systems.

(b) Surrogate variables. Surrogate variables can be sampled to provide information on other variables of primary interest. Specific conductivity and chlorophyll a are two examples of surrogate variables. Specific conductivity generally has a high correlation with total dissolved solids (TDS) concentrations. Measurement of TDS may be of major interest for irrigation purposes, for example, but is more expensive and time-consuming than the measurement of conductivity. This correlation should be determined for each reservoir since the correlation may vary among reservoirs. Certain phytoplankton species can cause taste and odor problems, clog water treatment plant filters, and create an unaesthetic appearance in the reservoir. Species enumeration and counting, however, can require special equipment and expertise that are not available in District or Division offices. Chlorophyll measurements can provide general information on the phytoplankton community but will not indicate the particular species affecting water quality.

(c) Specific variables. In situ and surrogate variables are appropriate for water quality surveys, but sampling programs designed to provide specific information for reservoir water quality management should supplement these variables with other constituents that relate to specific management objectives. At a minimum, in situ variables and Secchi depth should be measured at every sampling station on every sampling date, since these measurements add little to the cost but markedly increase knowledge about reservoir water quality. For example, sediment quantity and quality markedly influence reservoir water quality. Reservoir sedimentation surveys are periodically conducted to evaluate the loss of reservoir storage. An intensive water quality survey investigating longitudinal, lateral, and vertical water quality conditions should be conducted during July or August of the same year as the sedimentation survey. In situ variables, a representative nutrient, and chlorophyll should be measured, at a minimum, at selected locations along the sediment survey transects. These studies can provide a datum for comparisons among reservoir areas and among years.

(7) Sample handling and preservation. Appropriate sample handling and preservation is essential to ensure data quality. Standard Methods (APHA) (Ref. 48), EPA (Refs. 41, 43), Plumb (Ref. 20), and USGS (Ref. 46) discuss appropriate containers and proper preservation techniques for various water quality constituents and should be reviewed prior to field sampling (see also Table 5-3). Factors to be considered, in addition to those specifically mentioned in these references, include:

(a) Clean plastic containers are typically used for inorganic samples, with glass containers used for organic analyses. The caps or container lids, however, also must be of similar material to avoid sample contamination. Plastic lids or plastic liners and rubber stoppers can contaminate samples in glass bottles even though the surface area of the lid may be small.

(b) Proper sample preservation is critical if accurate and representative results are to be obtained from the sampling efforts. In general, all samples are placed on ice in the dark, even if additional preservation is required. Metal samples are generally preserved with nitric acid, nutrient samples with sulfuric acid, and organic samples through chilling.

(c) The desired form of a chemical species to be measured must be determined prior to preservation. For example, dissolved chemical species may require immediate field filtration prior to acidification. Acidification in the field, followed by laboratory filtration, can produce artificially high concentrations of dissolved elements. Appropriate filter pore sizes have been determined by APHA (Ref. 48).

(d) Analyses should be initiated as soon as possible after collection to avoid sample deterioration. Recommended and regulatory holding times are given in Table 5-3; these are based on Standard Methods (Ref. 48) and the USGS Handbook (Ref. 46).

(e) Any sample containing ≤ 0.5 mg/ℓ DO as measured onsite with a DO membrane electrode should be considered anaerobic. Anaerobic samples for those chemical parameters that may be either oxidized or precipitated on aeration can be pumped from the appropriate depths, through an acid-washed membrane filter if needed, into sample bottles containing the appropriate preservation. The sample bottom should be filled to capacity, then capped with an airtight lid.

(8) Sample identification. Proper sample identification in the field can eliminate subsequent problems in laboratory and statistical analyses. At a minimum, the date, reservoir station number, depth, preservation type, and replicate or split sample should be recorded on the sample. Before use in field sampling, all marking inks or fluids, labels, and containers should be tested under field conditions (i.e., immersion, agitation, preservation, spills) for container leakage, label and ink removal, and breakage. All previous labels or markings on sample containers should be removed prior to reuse. A separate log should be maintained that identifies the individual(s) collecting the samples, weather conditions, sample appearance, problems, unusual conditions, or other observations that may assist in interpreting the water quality data.

b. Field Quality Control and Assurance. Quality control (QC) and quality assurance (QA) must originate in the field if subsequent laboratory analyses are to provide accurate and representative data. Laboratory QA and QC programs have been established, but guidance on field QA and QC programs, although just as essential, is relatively new (Ref. 46). Field QA and QC can be improved through sampling standardization, sample preservation and replication, instrument calibration, and accurate records maintenance.

(1) Standardization. A standard approach to field sampling that incorporates checklists and a consistent sampling protocol can minimize omission and duplication errors during sample collection. This standard approach includes: a checklist of equipment and supplies required for each sampling trip; a specified location for the equipment and supplies in the sampling vehicle or boat; standard procedures for sample collection, filtration, and preservation; and routine procedures for delivering samples to the laboratory. One example of a standard procedure for sample preservation is color-coded labels for bottles (e.g., blue label for chilling, red for H_2SO_4 addition, and yellow for HNO_3 preservation). These codes can be combined (e.g., blue and red dot labels indicate chilling and H_2SO_4) or modified to indicate filtered versus nonfiltered samples. Standardization can also help ensure consistent results during personnel changes. Standard approaches, however, should not be confused with a rigid, inflexible program. Flexibility must be maintained to reflect changing program objectives and improved techniques and equipment. This flexibility also should allow for storm events and unusual field conditions or situations that may require collection of additional samples.

(2) Preservation and replication. A QC/QA program should quantitatively account for constituent additions or losses during sample preservation and determine the contribution of sampling error to the total sample error. Addition of known concentrations of reference standard solutions to field samples

should allow for measurement of error due to preservation, sorption, or interference. Sampling error also can be assessed through replication and split sampling. Replication involves collection of two or more distinct samples from the same location. Split sampling involves dividing a single sample among two or more sample containers for subsequent analysis. Estimates of sampling variability can be determined if as few as 10 to 15 percent of the samples receive known constituent additions and are replicated or split. Once the sources of error have been determined, sampling programs can be modified to minimize these errors.

(3) Instrument calibration. All instruments require periodic calibration. Changes in temperature, humidity, pressure, or other environmental factors can influence instrument calibration. In addition, transporting the instrument in a vehicle or boat can loosen sensitivity controls or connections and affect measurement accuracy. Calibration should be checked before and after each sampling trip, even for relatively stable probes such as temperature or specific conductance. Some instruments have internal temperature corrections, so temperature calibration can affect the accuracy of other constituent measurements. The precision and sensitivity of the instruments should be determined periodically for existing equipment and before initial field use for new equipment. Manufacturer's detection limits, precision, and sensitivity estimates are generally measured under ideal conditions and may not be applicable under field conditions. If two or more different manufacturers' instruments are used for measurements, they should be compared under field conditions even if the instruments were calibrated similarly. Instruments that cannot be calibrated should be used with caution. The manufacturers' manuals, Standard Methods (Ref. 48) and USGS Handbook (Ref. 46) should be consulted for specific recommendations on calibration.

(4) Records. Operating logs and records of measured field data, calibration curves, corrective actions, and QA activities should be maintained. Field personnel should be provided with a specified protocol for recording field observations, including content, format, names of the individuals collecting the data, and names of the individuals checking the validity of the data. These records should be identified and readily available for reference. Duplicate records may be a consideration. Guidelines should be established for record retention, duration, location, and assigned responsibility for each project. These guidelines and the retention period should be based on the sampling program design and objectives and the use of the data. Methodology generally changes significantly over a 10-year period and may influence data comparability over long time periods. Ten years may represent a minimum retention period for most monitoring and survey records.

c. Field Personnel. Since planning, engineering, and operational decisions that involve large expenditures of funds can be influenced by water quality data, these data must accurately represent the water quality conditions in the reservoir. The success of any sampling program ultimately depends upon competent laboratory and field personnel. This competence can be developed and maintained through on-the-job and formal training.

5-6. Laboratory Analysis. Laboratory analytical procedures and methodology are discussed in detail in EPA (Ref. 43), APHA Standard Methods (Ref. 48), and the USGS Handbook (Ref. 46). These references should be consulted for specific details on all procedures. Many laboratory analyses for Corps District and Division offices are contracted. Reference 42 discusses procedures for evaluating and monitoring laboratories. Engler (Ref. 65) provides guidance on contracting for laboratory analyses, while Peddicord (Ref. 19) provides guidance on contracting biological and chemical evaluations. Although the latter is directed toward dredged material, it discusses the contracting process, laboratory selection, contract management, and QA considerations. Contract management must be an active process since sample analysis represents the most important phase of the project. The effort put into contract management is directly proportional to the quality of the final product. The contract should have the flexibility to incorporate additional unscheduled sampling. Samples collected during the occurrence of storm events or other relatively rare situations in the field can provide the data and insight required for the development of better management approaches for reservoir water quality.

Section III. Database Management

5-7. Database Management Systems.

a. General. There are two major types of database management systems: general purpose or long-term databases, and specific water quality databases. General purpose databases such as STORET focus on information storage and retrieval strategies rather than on analysis of retrieved data sets. Specific water quality databases emphasize analytical and display routines. Short-term intensive studies or process-oriented studies typically require specific water quality databases.

b. General Purpose Database Management Systems. Water quality databases that are developed and used for general purposes strongly parallel the classical approach to database development and implementation strategies. General purpose database management systems focus primarily on database construction and retrieval strategies. Corps Divisions and Districts primarily use four data storage and retrieval systems: STORET, WATSTORE, AURAS, and SIR. Other systems, including UPGRADE, NAWDEX, and some that are commercially available, are also discussed in the following paragraphs.

(1) STORET, developed and operated by the EPA, maintains its water quality database on IBM equipment and is available to any user with the proper identification and access requirements. The EPA places few restrictions on its use and provides no automated QA controls over data entered into the system. Users inexperienced in the use of computers may have some difficulty in using the system. STORET has the capability of accessing the Statistical Analysis System (SAS) for statistical analyses.

(2) WATSTORE, developed and operated by the USGS, is available for use by Corps Division and District personnel on an AMDAHL computer, but the USGS is restrictive concerning potential users. Past users of WATSTORE expressed more confidence in data reliability than the users of STORET. The data in WATSTORE are verified and then transferred to permanent storage in WATSTORE and STORET. WATSTORE, like STORET, provides a limited number of statistical analysis packages, but it does provide the user with more advanced graphical techniques.

(3) AURAS, developed by the US Army Engineer Division, Ohio River, resides on the Computer Sciences Corporation INFONET system, which is on a UNIVAC 1108 that is being phased out. However, the AURAS program has been converted to Harris computers. AURAS provides the user with the ability to augment current Corps water quality databases but has limited statistical analysis and graphics or tabular display routines.

(4) The Scientific Information Retrieval System (SIR), developed by SIR, Inc., resides on the Boeing Computer Services CDC 7600 computer. SIR, a more advanced data storage and retrieval system than STORET, WATSTORE, and AURAS, interfaces the user with statistical analysis packages (Statistical Package for the Social Sciences and Biomedical Statistical Package) for analyzing extracted data. SIR does not maintain a nationwide water quality database; however, it does have the necessary language commands for building individual databases. SIR provides the user with good, reliable data storage and retrieval strategies along with a host of analytical capabilities.

(5) UPGRADE was developed and is maintained by the Council on Environmental Quality. The system resides on a commercial computer system and is accessible only by authorized users. UPGRADE has access to a resident water quality database and has an English language prompting command structure. UPGRADE has access to the SAS.

(6) NAWDEX was developed by the USGS and currently resides on the computer system at the USGS National Center in Reston, Virginia. NAWDEX primarily assists users in locating and retrieving information from other databases. Additional capabilities are not part of the software system. However, as a data storage and retrieval system, NAWDEX easily guides users through prompting sessions that allow the creation of or retrieval from the databases.

(7) Of the six systems discussed above, five are Government-sponsored database management systems. The remaining system, SIR, was developed by private enterprise and currently is available through Boeing Computer Services or may be adapted for use on in-house computer facilities. Other data management systems that are available for use include: Integrated Data Management System (IDMS), Information Management System (IMS), Adaptable Database Management System (ADABAS), or TOTAL. Database management systems are continually improving and becoming more user oriented. Periodic reviews of available systems are required.

c. Specific Water Quality Database Management and Analysis Systems. Specific water quality database management systems are characterized by their ability to simplify the database management storage and retrieval strategies. The commercial market for specific water quality database management systems is not as extensive as the general database management systems market. The five major software systems are described in the following paragraphs. All of these systems provide the user with a host of mathematical and statistical analysis routines. Each software package has advantages and disadvantages for particular applications and must be considered on a case-by-case basis.

(1) The Biomedical Statistical Package (BMDP) is marketed by BMDP Statistical Software, Los Angeles, California, and has provided statisticians, engineers, and scientists with a comprehensive and reliable software package for over 22 years. BMDP consists of 40 specialized statistical analysis routines, ranging from basic description statistics to a general linear models program that handles all aspects of unbalanced statistical designs. BMDP does provide for a limited amount of data storage and retrieval but, for large databases, it is better to use BMDP in conjunction with a front-end data storage and retrieval system. Graphics subsystems are not available with the complete BMDP package, while programming within the system is allowed on only a limited basis.

(2) The Statistical Package for the Social Sciences (SPSS), marketed by SPSS, Inc., of Chicago, Illinois, is available in two versions: the batch SPSS version and the conversation SPSS version known as SCSS. Both systems offer the user a wide selection of statistical analysis routines. However, only the SCSS version permits interactive communication between the database and the statistical analysis package. As with the BMDP, SPSS provides limited graphics capabilities, and if communication with a large, complex environmental database is required, a front-end data storage and retrieval system would be needed.

(3) The International Mathematical and Statistical Libraries (IMSL) is marketed by IMSL, Inc., Houston, Texas, and is probably the most complete package related to mathematical and statistical applications. IMSL is a FORTRAN-based system that consists of approximately 500 subroutines within the areas of mathematics and statistics. IMSL provides no interfacing programs between databases and the set of subroutines. A front-end storage and retrieval system, therefore, is mandatory. IMSL provides little or no graphics capabilities.

(4) Minitab, marketed by the Minitab Project at Pennsylvania State University, is the most recent research database management package marketed. Minitab has been customized to be compatible with a wide variety of computer systems, but currently has a limited amount of statistical analysis routines and graphics capabilities. The system is inexpensive and provides the user with easy access commands.

(5) The Statistical Analysis System (SAS), marketed by SAS Institute, Raleigh, North Carolina, is by far the most complete database management system currently marketed. SAS provides a wide range of mathematical and statistical analysis procedures, reliable database management strategies, color graphics, time series algorithms, report writing, operations research, and interface routines that permit interfacing between other mathematical and statistical packages and the IBM database management system known as the Information Management System (IMS). SAS is also programmable. This allows the user to construct mathematical algorithms not included in the basic package and incorporate them as library routines for later reference. SAS has an easy language command set and requires minimal training for the noncomputer-oriented individual. SAS is currently not available on the Harris system; it is available for the IBM and Digital systems.

5-8. Selection Criteria. System selection criteria should concentrate on the end-product user requirements and the available software packages. These software packages should provide the database system strategies, mathematical and statistical algorithms, graphic and tabular display routines, and programming capabilities, as well as compatability with other software packages so that other pertinent analysis packages can be incorporated easily. The system also should have reliable vendor support and be cost effective for the project. The database management system should focus on the needs of the engineer and scientist rather than on the requirements of the software system.

Section IV. Data Presentation

5-9. Methods. For all studies, data portrayal and display are a vital part of the data interpretation. The three methods of presentation most commonly used are: a complete listing of the database in some predetermined order, summary tables, and graphic displays. The complete database listing can be an important part of the data analysis but generally does not contribute significantly to data interpretation. Summary tables and graphic displays should be used; however, it is critical that these summary procedures enhance, not confuse, the water quality information presented.

5-10. Summary Tables. Summary tables should reduce the volume of data into a finite set of statistics, called descriptive statistics, which represent unbiased estimates of the unknown population parameters. Usually these tables consist of average or mean values and standard deviations of the population characteristics being investigated. Although the sample average and standard deviation are unbiased estimators, the complete picture of the underlying probability distribution that generated the measurements may not be adequately represented by only these two estimators. For example, if the probability distribution is symmetrical, then the sample mean, standard deviation, and confidence intervals adequately describe the underlying distribution. However, if the distribution is skewed, these estimators may be biased, and misinterpretations and erroneous conclusions can be drawn. To complement these estimators, such statistics as median, minimum, and maximum values and quartile points representing the 25th and 75th quartiles should also be displayed

(Table 5-4). With these additional statistics, the complete picture of the underlying probability distribution can be given.

5-11. Graphic Displays. Graphic displays complement summary tables. Graphic presentations can highlight areas that might not be detected from summary tables or rigorous statistical analysis of the data. A number of graphic displays that can be used to assess data distributions are discussed below.

a. X-Y Plots. Two-dimensional X-Y plots are the classical approach for graphically displaying the sample average and dispersion ($x \pm \sigma^2$) of population characteristics (Figure 5-2a). This plot displays two important characteristics: how the average value changes over sampling locations or time, and the homogeneity of variance assumptions, which are made in most poststudy analysis procedures. These plots can be improved by plotting not only the averages, but also the overall average with 100 $(1 - \alpha)$ percent confidence bands. (The letter α refers to the probability of error associated with the analytical procedure.) Presenting these features graphically can assist in data interpretation.

b. Quartile Plots. Quartile plots are an enhancement to the X-Y plots discussed above. This plot graphically displays the entire sample in the following manner. For the predetermined order, the maximum and minimum values, along with the 25th, 50th, and 75th quartile points, and the sample average can be plotted with a rectangular box enclosing the 75th and 25th quartile points. The interpretation of this plot can lead to discussion of symmetry and order statistics as they relate to differences among sampling locations or time. An example of this plot is shown in Figure 5-2b. As shown, the averages appear to be fairly uniform; however, skewness of the distributions is noticeable at locations B and C, whereas location A has a fairly symmetrical distribution. Furthermore, the medians (Q_{50}) are considerably different, and 50 percent of the data from location A and approximately 75 percent of the data from C are larger than the 75th quartile point for location B.

c. Tukey's Box Plot. Tukey's box plot resembles the quartile plot, but there are subtle differences between the two. The box plot, as with the quartile plot, will plot Q_{25}, Q_{50}, Q_{75}, and the sample average with a rectangle enclosing the 25th and 75th quartile points (see Figure 5-2c). The central vertical line extends from the box as far as the data extend or to a distance of, at most, 1.5 interquartile range; that is, the interquartile range is the distance between the 25th and 75th quartile points. Any value more extreme than this is marked with "0" if it is within 3 interquartile ranges of the box, or with an asterisk if it is still more extreme. Reference 105 provides more information about this plot.

d. Bar Charts. Bar charts (Figure 5-3a) are similar to X-Y plots, when averages are considered. The primary difference between the two is that X-Y plots depict the average as a single point while bar charts represent the average as a rectangular surface that encloses the area between the X-

TABLE 5-4

Example of Description Statistics That Can Be Applied to Water Quality Data

S T A T I S T I C A L   A N A L Y S I S   S Y S T E M

UNIVARIATE

VARIABLE POP   LABEL 1970 CENSUS POPULATION IN MILLIONS

MOMENTS

N	50	SUM WGTS	50
MEAN	4.0472	SUM	202.36
STD DEV	4.32932	VARIANCE	18.743
SKEWNESS	2.05522	KURTOSIS	4.54561
SS	1737.4	CSS	918.407
CV	106.971	STD MEAN	0.612258
T:MEAN=0	6.61028	PROB>|T|	0.0001
W:NORMAL	0.763044	PROB<W	0.01

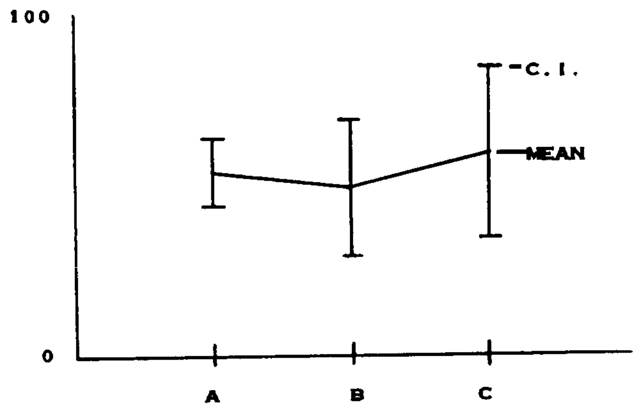
QUARTILES

100% MAX	19.95	99%	19.095
75% Q3	4.665	95%	11.495
50% MED	2.59	90%	10.65
25% Q1	0.97	10%	0.55
0% MIN	0.3	5%	0.385
		1%	0.3
RANGE	19.65		
Q3-Q1	3.695		
MODE	3.92		

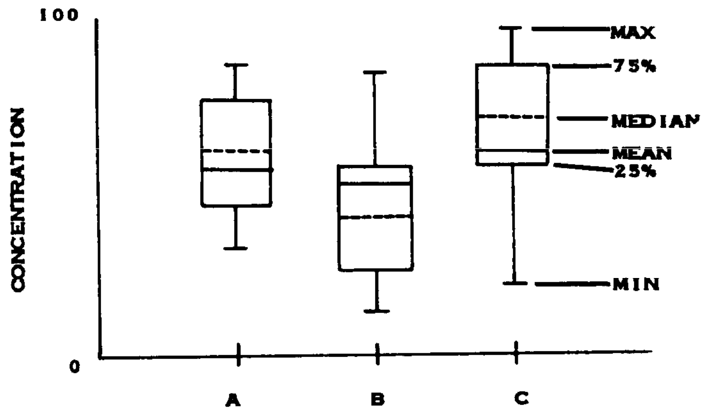
EXTREMES

LOWEST	ID	HIGHEST	ID
0.3	(ALASKA)	11.01	(ILL)
0.33	(WYO)	11.2	(TEXAS)
0.44	(VT)	11.79	(PA)
0.49	(NEV)	18.24	(NY)
0.55	(DEL)	19.95	(CALIF)

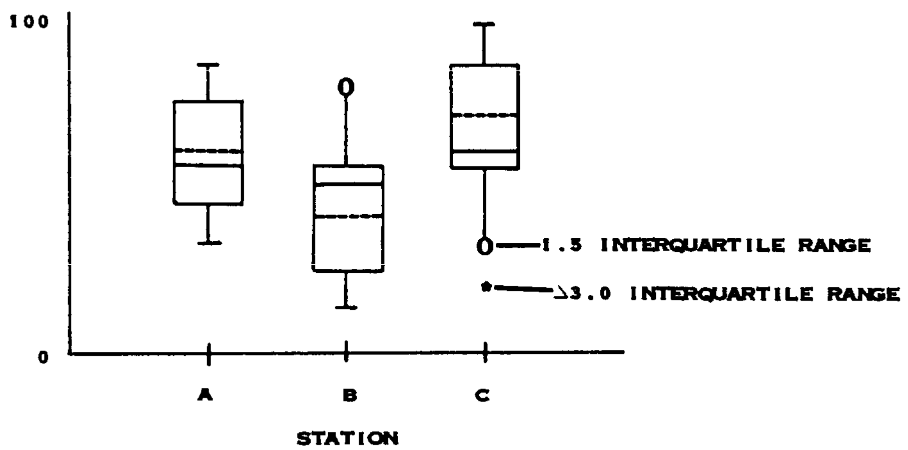
SOURCE: Ref. 96.



a. X-Y plots of mean and confidence intervals (C.I.)



b. Box plot with mean, median, and quartiles



c. Tukey's box plot with interquartile range

Figure 5-2. Example data plots

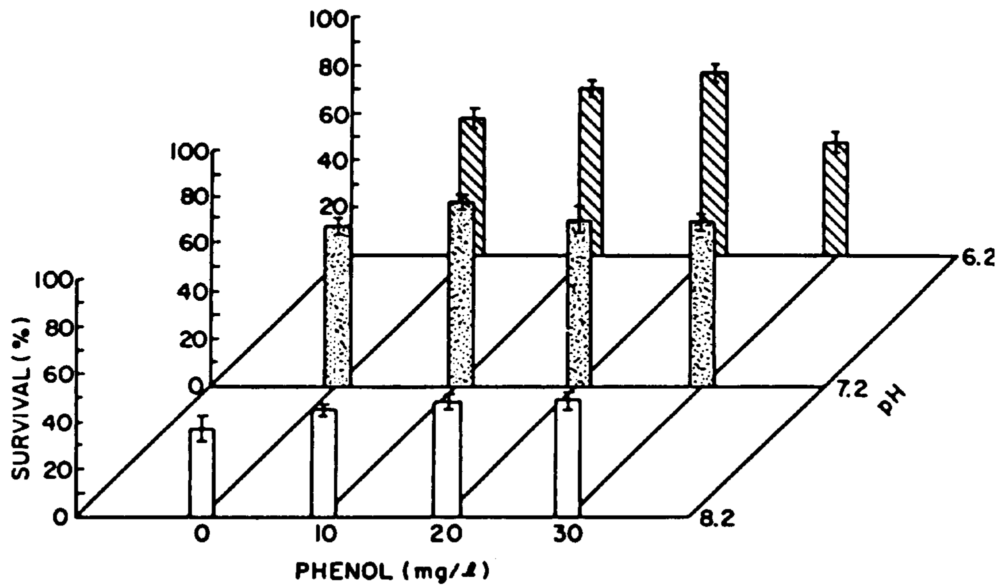
the average value. Although this figure may look very similar to a histogram, the two should not be confused, especially when the average value is being represented by the bar chart. Histograms have a continuous X-axis and the height of the bars represents frequency (Figure 5-3b). Although the bar chart is used primarily to plot average concentrations, it can and has been successfully used to display total values, frequency, percentages, etc. It is a flexible plot and can be easily extended to three dimensions.

e. Scatter Diagrams. Scatter diagrams are used to display relationships between two population characteristics, such as specific conductivity and time, TDS and specific conductivity, and total phosphorus and suspended solids (Figure 5-4). While these diagrams show trends and possible correlations, the appropriate statistical analysis should be performed prior to making inferences about the relationship between characteristics.

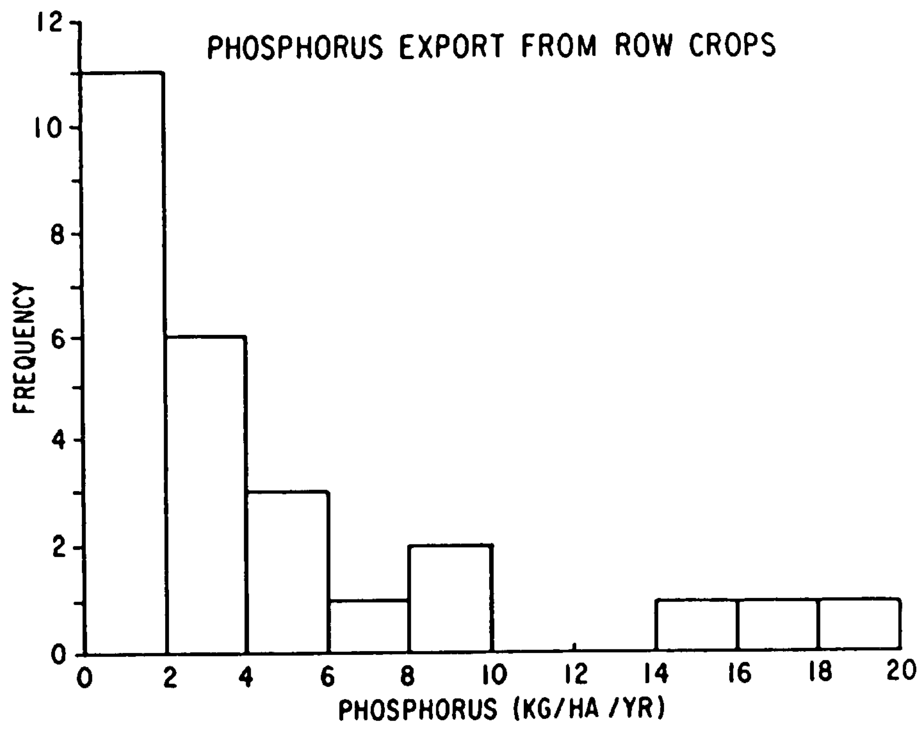
f. Annual Summary. At a minimum, inflow and outflow constituents should be plotted versus time on an annual basis. Water quality constituents collected at inpool stations should be plotted versus depth for each sampling date. Water quality constituents can be compared among years easily and quickly with graphic displays. Specialized annual summaries of reservoir constituents such as vertical (depth-time) and horizontal (depth-distance) isopleth diagrams can provide useful visual summaries for data presentation.

5-12. Quality Assurance. Quality assurance procedures are important not only for field sampling and laboratory analyses, but also for database management. Database entries should also be subjected to rigorous QA procedures so that aberrant and/or erroneous values can be removed or modified. Screening programs can be used to detect gross errors (e.g., transposition of pH value of 7.1 to 1.7); however, subtle errors such as pH 6.7 instead of 7.6 must be verified manually by comparing individual values. An error-free database is absolutely essential if valid scientific and statistical conclusions are to be derived.

5-13. Statistical Analysis. The use of appropriate statistical methods is essential for proper analysis and interpretation of reservoir water quality data and should be an integral part of the sampling program. Generally, this requires discussions with a statistician before the analyses are performed to minimize time and costs associated with the analyses. Statistical analyses should be performed only when all quality assurance checks have been satisfied, erroneous data entries have been removed, and the data have been graphically displayed. Statistical treatment of the data must be based on the sampling design and the assumptions made about the population characteristics under investigation. Statistical analyses are of two general types: parametric and nonparametric. Assumptions about the underlying probability distribution determine the appropriate statistical analyses. Before performing other statistical analyses, however, preliminary characteristics of the data can be determined by using descriptive statistics.

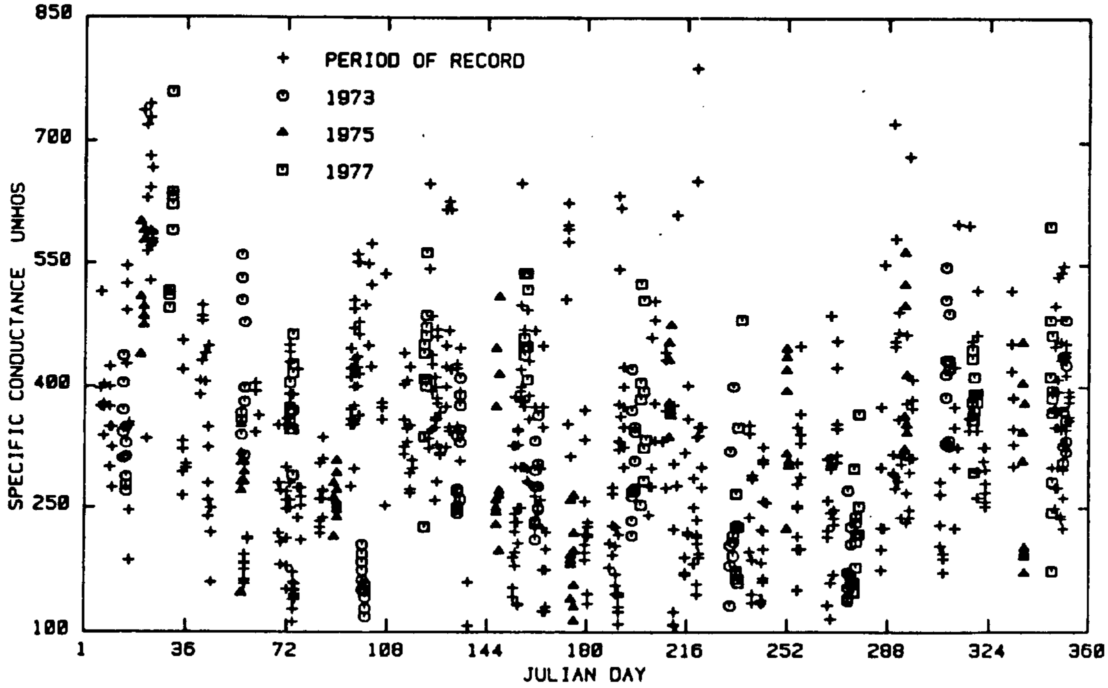


a. Sample bar chart

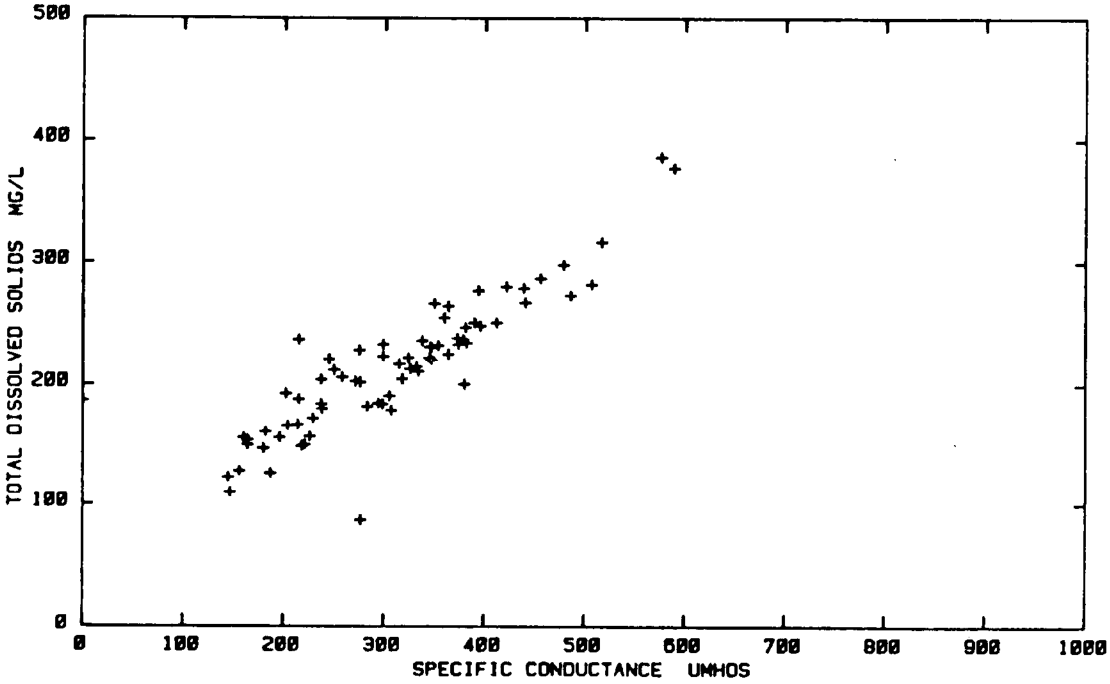


b. Sample histogram

Figure 5-3. Comparison of graphic display methods



a. Specific conductance versus time



b. TDS versus specific conductance

Figure 5-4. Example scatter diagrams

a. Descriptive Statistics. Descriptive statistics usually summarize or characterize a data set (Table 5-4). No assumptions about the probability distribution are made or implied for descriptive statistics. The characteristics may be a mean, median, mode, variance, range, etc. Each estimator summarizes a data set in a unique way and represents a specific population characteristic. For example, the mean is the arithmetic average of n sample values, while the median represents the middle value of n measurements. The mode is defined as the measurement with the maximum frequency. These characteristics provide information on the sample population. The mean and median are similar, for example, if the distribution of values is symmetrical but will diverge as the distribution becomes skewed. The mean is more influenced by outliers than the median. Quantitative measures of the data distribution such as the range, variance, or standard deviation provide information on the dispersion of the data. Descriptive statistics should be computed as part of the initial data evaluation. Additional information can be found in Snedecor and Cochran (Ref. 99) or Steel and Torrie (Ref. 101).

b. Parametric Statistics. Parametric statistical inferences about a finite number of unknown population parameters are based on an underlying density or probability function such as the normal distribution, binomial distribution, Poisson distribution, etc. Since all measurements are subject to error, the usual assumption is that these measurement errors have an underlying normal distribution with a zero mean and an unknown variance. The error terms are assumed to be independent, identically distributed, generally with a normal distribution (i.e., homogeneity of variances), and additive. The significance of these assumptions is discussed in Snedecor and Cochran (Ref. 99), Sokal and Rohlf (Ref. 100), and Steel and Torrie (Ref. 101). These assumptions can be represented by the linear model

$$Y_i = \mu + \varepsilon_i$$

where

Y_i = measurement made on the i^{th} sample ($i = 1,2...n$)

μ = population mean of the measurements

ε_i = the error involved in making the i^{th} measurement (i.e., the deviation of the i^{th} measurement from the unknown population mean)

With this as the underlying model and the above assumptions, the measurement Y_i will possess a normal distribution with mean μ and variance σ^2.

(1) Statistical inference. When sample characteristics are used to infer some information about the general population, the subject is called inductive statistics or statistical inference. Inference becomes a scientific method, differentiating it from mere guessing, when probability statements concerning the accuracy of the estimate or reliability of a decision are

incorporated. The two types of problems most frequently encountered are estimation and test of hypotheses.

(a) Estimation. The area of estimation considers questions such as: what are the estimates of the unknown population parameters based on the sample measurements, and what properties do these estimators exhibit? The class of estimators may be divided into two areas: point estimates and interval estimates. The usual estimators are the sample average and variance, which estimate the population mean and variance, respectively, and can either be reported as point estimates or interval estimates. The interval estimate for the mean is the 100 $(1 - \alpha)$ percent Student confidence interval, while the 100 $(1 - \alpha)$ percent confidence interval for the population variance is based on the Chi-square distribution. The symbol α represents the probability of rejecting the hypothesis when this hypothesis is true in favor of an alternative hypothesis that is false. This level has typically been $\alpha = 0.05$ but can be set at any level, i.e., 0.20 or 0.01. Interval estimates can be extended to the two-sample case, i.e., paired samples or independent samples. (Further information is given in Refs. 99 and 101.) Population parameters can be estimated in three ways: maximum likelihood techniques, least squares techniques, and the method of moments. With the assumption of normality, these estimation methods produce similar results. However, if the normality assumption is violated, the estimates may vary considerably. In order to determine which estimator is best, three properties of estimators should be considered: bias, consistency, and minimum variance. For more information on these techniques and properties, see Mood and Graybill (Ref. 88), Hogg and Craig (Ref. 76), or Kendall and Stuart (Ref. 83).

(b) Test of hypotheses. The test of hypotheses is an extension of interval estimation. Confidence intervals usually concentrate on one or two sample procedures, while hypothesis testing extends this concept to multiple samples. The underlying goal of hypothesis testing is to assist in making decisions about differences among population characteristics. Statistical analysis only assists in the decision process, however. Statistics is a tool that must be used with good engineering and scientific judgment in making decisions about population characteristics. One- and two-sample tests, the most elementary form of hypothesis testing, involve decisionmaking about population parameters that characterize either one or two populations. Decisions can be made on a population mean, differences in population means, effects of treatments on a population mean, a population variance, or homogeneity of population variances. Statistical tests appropriate for decisionmaking about each of these areas are: the one-sample t-test, the two-sample independent t-test, the paired t-test, the Chi-square test, and the F-test, respectively. Several of these tests are summarized in Table 5-5, and explanations are given in Snedecor and Cochran (Ref. 99), Sokal and Rohf (Ref. 100), and Steel and Torrie (Ref. 101). When the statistical problem involves two or more factors, the appropriate statistical test is the analysis of variance (ANOVA). Essentially, the ANOVA partitions the total variance into known sources of variation. From here, significant and nonsignificant contributions to the total variance can be observed. For example, in investigating the effects of a

TABLE 5-5

Summary of Parametric Statistical Tests

One-Sample	Two-Sample		k-Sample	Measures of Association
	Paired Samples	Independent Sample		
A. Student's t-test	A. Paired t-test	A. Independent E-test	A. One-way ANOVA	A. Linear Correlation
Assumptions:	Assumptions:	Assumptions:	B. Randomized Complete Block	B. Linear Regression
1. Normal distribution	1. Normality	1. Normality	C. Factorials	C. Nonlinear Regression
2. Unknown variance	2. Unknown variance	2. Unknown, but common variances	D. Nested Designs	D. Time Series
	KEY: Two experiments performed on one sampling unit.		E. Nested Factorials	E. Analysis of Covariance
		KEY: Independent sampling units.	F. Randomized Incomplete Block Design	F. Response Surface Models
	3. Hypothesis	3. Hypothesis	Assumptions:	
	$H_o: \mu = 0$	$H_o: \mu_1 = \mu_2$	1. Normality	
		B. F-test	2. Equal variances	
		Assumptions:	3. Additive error term	
		Normality	Transformations:	
		KEY: Test hypothesis of equal variances.	1. Logarithmic	
			2. Square root	
			3. Arcsine	

hydropower reservoir on a river, the design could be to establish river stations above and below the hydropower reservoir in the main channel and along the banks and to collect both surface and bottom samples. The sampling units are subjected to outside sources of variations. Such sources could be treatment or factor effects due to the lack of homogeneity among sampling units. These outside sources of variation can be identified and explained by partitioning the total variance into the variance attributable to each of these outside influences, such as main channel and bank locations or above- and below-reservoir responses. An F-test, which is a test for the equality of two variances, can be constructed so that significance or nonsignificance of each of these factors can be determined. For the significant effects, post-ANOVA procedures can be used to separate the treatment effects such as main channel versus bank locations. More detailed information concerning the ANOVA, post-ANOVA tests, and transformation of data to satisfy the normality assumption is given in Refs. 99 and 101. General characteristics are summarized in Table 5-5.

(2) Regression analysis.

(a) Regression analysis is divided into two categories: cause-and-effect models and correlation and regression models. The primary distinction between these two categories is the assumptions placed on the measured variables. With cause-and-effect models, the relationships between a dependent response variable and a set of independent variables are studied. These models assume the independent variable (i.e., X variable) is fixed or measured without error. This means only the dependent variable (i.e., Y variable) is a random variable and includes random error. Usually, the models are defined as a polynomial relationship between the independent and dependent variables. Multiple linear relationships also can be defined, as long as the assumption of fixed independent variables is preserved. Some typical cause-and-effect relations are:

$$Y = \sum_{i=0}^{p} \alpha_i X^i \qquad \text{(general polynomial model)}$$

$$Y = \exp\left(\sum_{i=0}^{p} \alpha_i X^i\right) \qquad \text{(general exponential model)}$$

$$Y = \sum_{i=0}^{p} \alpha_i X_i \qquad \text{(general multiple linear model)}$$

where

α_i = unknown parameters

X^i = an independent variable raised to the i^{th} power

X_i = the i^{th} independent variable

(b) The estimation technique used for providing estimators of the unknown α_i values is the least squares technique. Correlation and regression models do not adhere to the assumption of fixed independent variables. With this type of condition, several characteristics about the population are measured from the same sample unit. For example, assume temperature, suspended solids, DO, Fe, Mn, and H_2O measurements are made on water samples. The aim is to investigate the multivariate relationship among these variables. The procedure, which provides an interdependent structure among these measurements, is correlation and regression. The regression model is usually a multiple linear model, and the regression parameters α_i are estimated by the least squares technique. The subtle differences between these two types of regression problems are the estimation of error variance and the measure of goodness of fit. With the cause-and-effect model, the error variance estimate can be obtained from the ANOVA table while partial correlation and variance estimates have to be computed separately with the correlation and regression model. With the cause-and-effect model, the coefficient of determination R^2 is used to discuss the percent of the total variance explained by the model, while the sample correlation coefficients and partial correlation coefficients are used to discuss significant relationships in correlation and regression models. Other measures of association are listed in Table 5-5. A more complete discussion of this subject is found in Drapier and Smith (Ref. 63), Snedecor and Cochran (Ref. 99), and Steel and Torrie (Ref. 101).

(3) Time-series. Observations on a population characteristic through time generate an ordered set of data known as a time series. The values assumed by a variable at time t may or may not embody an element of random variation but, in a majority of the cases, random variation will be present. Time-series analysis can be used to investigate: a trend, or long-term movement; oscillations about the trend of greater or less regularity; a seasonal effect; and a random, unsystematic, or irregular component. Time-series analysis assumes data were collected at regular intervals (i.e., daily, weekly, or 30-day) and generally requires relatively long records. A more complete discussion of time-series analysis can be found in Box and Jenkins (Ref. 54) or Kendall and Stuart (Ref. 83).

c. Nonparametric Statistics. The classical approach to data analysis is parametric statistics. However, there are alternative procedures called distribution-free inference, which are as reliable and robust as the classical parametric procedures. The complete area of distribution-free inference is known as nonparametric statistics. In a distribution-free inference, for testing or estimation, assumptions regarding the specific underlying population distribution are not necessary. The term nonparametric test implies a test for a hypothesis that is not a statement about population parameter values, such as the mean and variance. The type of hypothesis, then, considers only the form of the population, as in goodness-of-fit tests, or some characteristic of the probability distribution of the sampled data, as in tests for randomness and trend. Distribution-free test and nonparametric test are not synonymous labels, since distribution-free test relates to the distribution of the test statistic while the nonparametric test refers to the type of hypothesis being tested. Many parametric tests have nonparametric or distribution-free equivalences (Table 5-6). A complete discussion may be found in Gibbons (Ref. 68) and Siegel (Ref. 98). The following paragraphs introduce some of these procedures.

(1) Tests on goodness of fit. An important problem in statistics relates to obtaining information about the form of the population from which the sample is drawn. For example, the traditional parametric test, based on Student's t-distribution, is derived under the assumption of a normal population. The exact distribution theory and probabilities of making Type I and Type II errors depend on this population form (i.e., Type I = rejecting a hypothesis when it is true; Type II = accepting a hypothesis when it is false). Therefore, it might be desirable to check the reasonableness of the normal assumption before forming any conclusions based on the t-distribution. Tests based on the distribution inference are called goodness-of-fit (GOF) tests. Since a Chi-square distribution can be generated from a population of standard normal deviates, a Chi-square test can be used to compare the cumulative distribution of the sample population with a Chi-square distribution. The Kolmogorov-Smirnov one-sample test is another nonparametric test that is applicable to continuous frequency distributions. In many cases, it has greater power than the Chi-square test for GOF. Additional information can be found in Gibbons (Ref. 68), Snedecor and Cochran (Ref. 99), and Sokal and Rohlf (Ref. 100).

(2) General two-sample problem. The analogous parametric problem to the nonparametric two-sample problem is the independent t-test. The t-test checks for equality of means under the assumption of normality and homogeneous variances while the nonparametric tests emphasize differences in location, scale, and medians. Some of the more common tests are the Wald-Wolfowitz runs test, the Kolmorgorov-Smirnov two-sample test, the median test, and the Mann-Whitney U-test (see Refs. 68 and 98; Table 5-6).

(3) Test of equality of k independent samples. Extensions of the two-sample tests are available in nonparametric statistics similar to procedures. Some applicable nonparametric tests for comparing k-samples are: the k-sample median test, the Kruskal-Wallis one-way ANOVA test, the Friedman two-way ANOVA test by ranks, and Conover's k-sample slippage test (Ref. 60).

(4) Measures of association for bivariate samples. There are equivalent nonparametric tests to measure the association between variables analogous to

TABLE 5-6

Summary of Nonparametric Statistical Tests

Level of Measurement	One-Sample Case	Two-Sample Case: Related Samples	Two-Sample Case: Independent Samples	k-Sample Case: Related Samples	k-Sample Case: Independent Samples	Measure of Correlation
Nominal[1]	Binomial test Chi-square test	McNemar test for the significance of changes	Fisher exact probability test Chi-square test for two independent samples	Cochran Q test	Chi-square test for k independent samples	Contingency coefficient
Ordinal[2]	Kolmogorov-Smirnov one-sample test One-sample runs test	Sign test Wilcoxon matched pairs signed-rank	Median test Mann-Whitney U Test Kolmogorov-Smirnoff two-sample test Wald-Wolfowitz runs test	Friedman two-way analysis of variance	Extension of median test Kruskal-Wallis one-way analysis of variance	Spearman rank correlation Kendall rank correlation Kendall partial rank correlation coefficient Kendall coefficient of concordance

[1]When numbers of other symbols are used to classify a characteristic of a population, but the order of magnitude has no bearing on the outcomes of the classification scheme: arbitrary scale.

[2]When numbers or other symbols are used to classify a characteristic of a population and the order of magnitude implies a mathematical relationship.

parametric measures of association or correlation. These measures are Kendall's tau coefficient and Spearman's coefficient of rank correlation (Table 5-6). These measures of association are as powerful as the classical approach to correlation (e.g., the Pearson product-moment correlation coefficient) and are more easily computed. Gibbons (Ref. 68) and Siegel (Ref. 98) provide information on these techniques.

d. Discussion. This section has briefly discussed some essential components of statistical analysis, including both parametric and nonparametric analyses. Parametric statistics rely on certain assumptions about the underlying probability distribution and concentrate in the areas of estimation and hypothesis testing about the unknown population parameters. Nonparametric procedures make no assumptions about the distributional properties of the population and allow for broader hypotheses or inferences to be examined. Each area, parametric and nonparametric, can provide the engineer and scientist with valid and reliable results. Within each area, the appropriate statistical tests are ultimately based on the study objectives and sampling program design.